

IMAGE PROCESSING METHOD AND DEVICE THEREFOR

Publication number: JP9167233

Publication date: 1997-06-24

Inventor: ISHITANI YASUTO; UDA AKIHIRO

Applicant: TOKYO SHIBAURA ELECTRIC CO

Classification:

- international: G06K9/20; G06K9/62; G06T7/00; G06T7/40; G06K9/20;
G06K9/20; G06K9/62; G06T7/00; G06T7/40; G06K9/20;
(IPC1-7): G06T7/00; G06K9/20

- European:

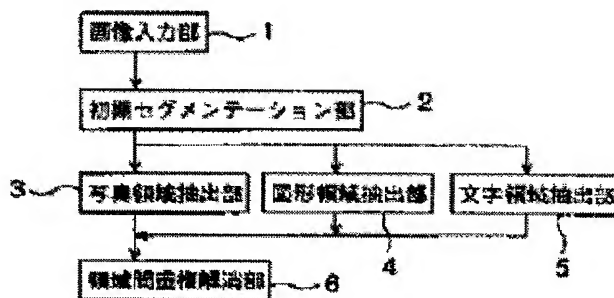
Application number: JP19950327261 19951215

Priority number(s): JP19950327261 19951215

Report a data error here

Abstract of JP9167233

PROBLEM TO BE SOLVED: To discriminate with high accuracy plural image areas having different properties of the character, photo and graphic areas, etc., from each other, in a document image that has its unknown format and unknown document input direction respectively. **SOLUTION:** The black connection component is extracted out of the document image inputted via an image input part 1 and assigned to a photo area extraction part 3, a graphic area extraction part 4 or a character area extraction area 5 according to the circumscribed rectangle size of the extracted component. Each of these areas 3 to 5 extracts an image area accordant with its property, and if the spatial overlapping is detected with its property, and if the spatial overlapping is detected among the areas extracted by the parts 3 to 5, an area overlapping dissolution part 6 extracts a combination based on the exclusive and coexistent relating among the areas 3 to 5.



(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-167233

(43)公開日 平成9年(1997)6月24日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 7/00			G 0 6 F 15/70	3 3 0 Q
G 0 6 K 9/20	3 4 0		G 0 6 K 9/20	3 4 0 L

審査請求 未請求 請求項の数8 O L (全 15 頁)

(21)出願番号 特願平7-327261

(22)出願日 平成7年(1995)12月15日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 石谷 康人

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 宇田 明弘

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

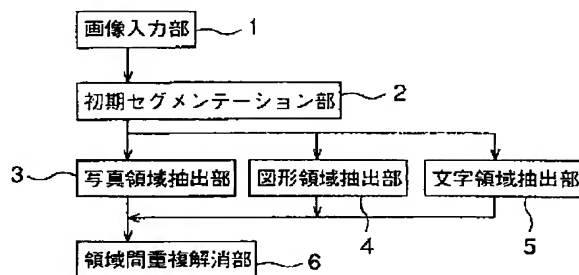
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 画像処理方法および画像処理装置

(57)【要約】

【課題】書式が未知であり、文書の入力方向が未知である文書画像において、文字領域、写真領域、図形領域等の性質の異なる複数の画像領域を高精度に識別できる画像処理方法および画像処理装置を提供できる。

【解決手段】画像入力部1で入力された文書画像から黒連結成分を抽出して、その外接矩形のサイズから写真領域抽出部3、図形領域抽出部4、文字領域抽出部5のいずれかに振り分け、それぞれにおいて、各画像領域の特性に合った画像領域抽出処理を行い、写真領域抽出部3、図形領域抽出部4、文字領域抽出部5のそれぞれで抽出された領域間に空間的重なりが見られる場合には、領域重複解消部6で各領域間の排他関係と共存関係に基づき、組み合わせを抽出する。



【特許請求の範囲】

【請求項 1】 入力された文書画像から性質の異なる画像領域を識別する画像処理方法において、前記文書画像から前記画像領域の性質に応じて部分領域の抽出を行い、この抽出された画像領域間の排他的関係と共存関係に基づいて、前記抽出された複数の画像領域間の重複領域が属する画像領域を識別することを特徴とする画像処理方法。

【請求項 2】 入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出し、この部分領域から抽出された文字列から文字を切り出して文字認識し、この認識結果に基づく評価値が最良となるまで、前記統合条件を変更しながら前記抽出された部分領域を統合することにより文字領域を抽出することを特徴とする画像処理方法。

【請求項 3】 入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出し、この部分領域から抽出された文字列の幾何的情報を抽出して、この幾何的情報を基に前記部分領域の評価値を算出し、この評価値が最良となるまで、前記統合条件を変更しながら前記抽出された部分領域を統合することにより文字領域を抽出することを特徴とする画像処理方法。

【請求項 4】 入力された文書画像から抽出された文字成分の隣接関係を抽出し、この抽出された文字成分の隣接関係に基づき前記文字成分をグループ化して部分領域を生成し、この生成された部分領域を構成する文字成分とその隣接関係に基づき文字領域を抽出することを特徴とする画像処理方法。

【請求項 5】 入力された文書画像から性質の異なる画像領域を識別する画像処理装置において、前記文書画像から前記画像領域の性質に応じて部分領域の抽出を行う部分領域抽出手段と、この部分領域抽出手段で抽出された部分領域間の排他的関係と共存関係を管理する管理手段と、この管理手段で管理された部分領域間の排他的関係と共存関係に基づいて、前記部分領域抽出手段で抽出された複数の部分領域間の重複領域が属する画像領域を識別する識別手段と、を具備することを特徴とする画像処理装置。

【請求項 6】 入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出する抽出手段と、この抽出手段で抽出された部分領域から文字列を抽出する文字列抽出手段と、この文字列抽出手段で抽出された文字列から文字を切り出して認識する文字認識手段とを具備し、この文字認識手段での認識結果に基づく評価値が最良となるまで、前記統合条件を変更しながら前記抽出手段で抽出された部分領域を統合することにより、文字領域を抽出する画像処理装置。

【請求項 7】 入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出する抽出手段と、

この抽出手段で抽出された部分領域から文字列を抽出する文字列抽出手段と、

この文字列抽出手段で抽出された文字列の幾何的情報を抽出する手段と、

前記抽出された文字列の幾何的情報を基に前記部分領域の評価値を算出する手段とを具備し、

10 前記評価値が最良となるまで、前記統合条件を変更しながら前記抽出手段で抽出された他の部分領域を統合することにより、文字領域を抽出することを特徴とする画像処理装置。

【請求項 8】 入力された文書画像から抽出された文字成分の隣接関係を抽出する第 1 の抽出手段と、

この第 1 の抽出手段で抽出された文字成分の隣接関係に基づき前記文字成分をグループ化して部分領域を生成する生成手段と、

20 この生成手段で生成された部分領域を構成する文字成分とその隣接関係に基づき文字領域を抽出する第 2 の抽出手段と、

を具備したことを特徴とする画像処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、文字、写真、絵、図形、表などが混在する文書画像からそれぞれの領域を識別する画像処理方法、および、それを用いた画像処理装置に関する。

【0002】

30 【従来の技術】最近、書類形態として蓄積されている大量の情報を計算機に自動入力できるシステムの実現に対する要求が非常に高まっている。このようなシステムを実現する場合には、文書画像をデジタル画像として計算機に取り込み、文書の意味的、幾何的性質から、文字領域、写真領域、図形領域などの質の異なる領域を自動的に分離する機能（領域分割）が重要となる。

【0003】この機能により得られた各々の領域ではその性質に基づいた認識処理が施されることによって利用価値の高い情報を計算機に入力でき、活用することが可能となる。

【0004】これまでに提案されている領域分割方式は、例えば、「Proc. 12th ICPR, pp 345-349, (1994)」、「信学論 D-11, Vol. 1, J75-D-11, No. 2, pp 246-256, (1992)」に記載されているように、1種類の幾何特徴を文書画像で抽出し、その幾何的性質および分布により上記 3つの領域に分離するものであった。

50 【0005】この方式では、処理に用いた幾何特徴が文字、写真、図形の各々の領域の幾何的性質を適切に表現しているものでなければならないが、現状ではそのよう

なものは見つかっておらず、従って十分な分離能力が得られていない。

【0006】さらに、「Proc. 1st ICDA R, pp945-962, (1991)」、「信学論D-11, Vol. J72-D-11, No. 1, pp93-104, (1989)」に記載されているように、文書の背景部(空白領域)に着目して文書画像を分割する方法がある。

【0007】この方式では、オブジェクトが密に分布している文書の空白領域とスペースに分布している文書の空白領域を明確に区別することができず、従って、異なる領域をまとめて一つの領域として抽出してしまったり、同質な領域を分離してしまうという欠点がある。

【0008】また、「信学論D-11, Vol. J78-D-11, No. 3, pp465-473, (1995)」、「Machine Vision and Applications, Vol. 7, pp. 237-246, (1994)」、「IEEE Trans. Pattern Analysis and Machine Intelligence Vol. 15, No. 11, pp. 1162-1173 (1993)」に記載されているように、文字成分にのみ着目して文字列とその塊であるブロックを抽出し、それ以外を非文字領域として無視する手法がある。

【0009】この方式では、

- ・ 文字らしきものを文書画像から抽出しそれらを順次統合していった、整列している文字列のみを抽出するが、そのような条件が非文章領域でも満たされる場合にはその領域を文字領域とみなしてしまう。

- ・ 規則的な整列性を重視する余り、「文字が不規則に分布している文字列」を正しく抽出することはできない。

- ・ 文字成分のみに着目したアプローチでは写真領域と図形領域を識別することはできない。

- ・ 縦書き文章と横書き文章が混在する文書を処理することはできない。

といった問題点がある。

【0010】

【発明が解決しようとする課題】前述の従来方式の何れも単一な幾何特徴のみに基づいて文書画像を文字領域、写真領域、図形領域に分類するので、分類処理に必要とされる情報が十分に得られず高精度な処理結果を得ることができないという問題点があった。

【0011】また、文字領域においては、種々の文字サイズ・字間・行間を持つ文字列が混在する文字が二次元的に配置されている(縦書き文字列と横書き文字列が混在する)文書から各々の文字列を高精度に抽出することはできないという問題点があった。

【0012】また、図形領域中の文字列は図形扱いとしようために、検索などにおいて有効に活用すること

はできなかった。さらには、文書画像は必ず正しい方向で入力されることを前提としているため文書を入力する際に文書の方向に関して細心の注意を払わなければならないという問題点があった。

【0013】そこで、本発明は上記問題点に鑑みてなされたものであり、書式が未知であり、文書の入力方向が未知である文書画像において、文字領域、写真領域、図形領域を高精度に分離・識別でき、さらには、横書き文章と縦書き文章を分離して抽出できる画像処理方法および画像処理装置を提供することを目的とする。

【0014】

【課題を解決するための手段】本発明の文書認識方法は、入力された文書画像から性質の異なる画像領域を識別する画像処理方法において、前記文書画像から前記画像領域の性質に応じて部分領域の抽出を行い、この抽出された画像領域間の排他的関係と共存関係に基づいて、前記抽出された複数の画像領域間の重複領域が属する画像領域を識別することにより、文書画像を構成する文字成分・写真成分・図形成分の各々に対応した領域抽出手段を備えており、領域間で重複が生じてもそれを解消できるので、文字領域と写真領域と図形領域を高精度に識別・分類することができる。また、それぞれの領域に対応した抽出処理を行うことにより、領域分割の性能が各々の領域の文書中に占める割合に依存しない。

【0015】また、本発明の画像処理装置は、入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出し、この部分領域から抽出された文字列から文字を切り出して文字認識し、この認識結果に基づく評価値が最良となるまで、前記統合条件を変更しながら前記抽出された部分領域を統合することにより文字領域を抽出することを特徴とする。

【0016】また、本発明の画像処理装置は、入力された文書画像から抽出された文字成分に基づく統合条件に従って部分領域を抽出し、この部分領域から抽出された文字列の幾何的情報を抽出して、この幾何的情報を基に前記部分領域の評価値を算出し、この評価値が最良となるまで、前記統合条件を変更しながら前記抽出された部分領域を統合することにより文字領域を抽出することを特徴とする。

【0017】さらに、本発明の画像処理方法は、入力された文書画像から抽出された文字成分の隣接関係を抽出し、この抽出された文字成分の隣接関係に基づき前記文字成分をグループ化して部分領域を生成し、この生成された部分領域を構成する文字成分とその隣接関係に基づき文字領域を抽出することを特徴とする。

【0018】このような特徴により、文字領域を抽出する際、文字列の傾き方向を検出し、縦書き／横書きに対応して文字列を抽出し、得られた文字列を上下左右の4方向で文字認識して、その文字認識結果の有効性を評価するので、入力文書が

- ・縦書き文章と横書き文章が混在している場合
 - ・文書が傾いている場合(部分文章領域で独立して傾いている場合を含む)
 - ・文書の入力方向が未知である場合
- を取り扱うことができる。

【0019】このように領域分割処理に文字認識結果を利用すると、

- ・文章領域と非文章領域を高精度に識別することが可能となる

- ・図形領域中の文字成分を抽出することができる
- という利点もある。

【0020】また、本発明では、複数の領域分割処理結果を出力することができ、これをオペレータに提示し、オペレータがその中から最良のものを選ぶというインタラクティブで簡便な作業を実現できる。この結果、常に高精度な出力結果を得ることが可能である。

【0021】

【発明の実施の形態】本発明の実施形態について、図面を参照して説明する。まず、用語について説明する。以下の説明において、「画像オブジェクト」とは、画像中の実体、すなわち、例えば2値画像の場合、実体とは黒画素の任意の集合のことをいう場合がある。

【0022】また、「部分領域」とは、一つあるいは複数の画像オブジェクトによって構成されるある大きさを持つ2次元領域のことをいう場合がある。図1は、本発明の一実施形態に係る画像処理装置の構成を概略的に示したものである。

【0023】画像入力部1から例えば光学的に取り込まれた文書画像は、公知である2値化処理手法により白と黒の2値の画像データに変換され、この2値化画像について例えば文献「信学技報、PRU92-32、1992」に記載された傾き検出・補正処理を施し、傾きのない2値画像に変換する。以後、この画像を入力画像とする。

【0024】次に、入力画像は初期セグメンテーション部2において、初期セグメンテーション処理により画像オブジェクトを写真、図形、文字のいずれかに分類する。そして分類された画像オブジェクトに対して、それぞれ、写真領域抽出部3、図形領域抽出部4、文字領域抽出部5において、部分領域を抽出して文字領域、図形領域、写真領域をそれぞれ抽出する。

【0025】抽出結果において解釈の曖昧さが見られる(例えば、部分領域に重なりが生じている場合など)には、領域重複解消部6で文書画像全体あるいは部分領域間の整合性や各部分領域の秩序性などを調べることでより曖昧さを解消して高精度かつ信頼性の高い処理結果を得ることができる。

【0026】次に、初期セグメンテーション部2の処理について説明する。まず、入力画像に対し公知のラベリング処理を行って、黒連結成分を抽出する。得られた連

結成分をその外接矩形で囲み、得られた外接矩形のサイズ(横幅: w と縦幅: h) に基づいて、

- ・文字候補矩形 (w と h のどちらかがしきい値 th_{s1} を満たす矩形)

- ・ドット候補矩形 (w と h の両方がしきい値 th_{s2} を満たす矩形、微小矩形とも呼ぶ)

- ・その他(上記1、2以外の矩形、巨大矩形とも呼ぶ)に分類する。

【0027】ここで、 th_{s1} 、 th_{s2} を予め設定した値とする。次に、写真領域抽出部3の処理について説明する。2値画像では写真領域は「べた塗り領域」かあるいは「ソルト・ペッパー状のドット領域」と組み合わせられた形で出現することが多いことからべた塗り領域とドット領域をそれぞれ抽出し、それらのうち近接するものをまとめて写真領域と判別する。

【0028】まずドット領域の抽出処理について説明する。初期セグメンテーション処理によりドット候補矩形に分類されたものに対して、近接するものをまとめていき、幾つかのクラスタ(部分領域)を作る。このうちクラスタの大きさ(例えばクラスタを外接する矩形を検出しその縦幅および横幅の両方)が予め定めたしきい値 th 。以下のものをノイズと判定してもよい。得られたクラスタの集合を $\{Dots\}$ と表す。 $\{Dots\}$ の各要素はそれらを内接する多角形で存在範囲が示されている。

【0029】べた塗り領域の抽出処理について説明する。初期セグメンテーション処理で「巨大矩形」と判定されたものに対して、水平方向および垂直方向に長さ th_1 (予め定めたしきい値) 以上の黒ランを抽出し、それらに対してさらに上記ラベリング処理を行う。得られた黒連結成分を多角形で内接し、その多角形内で黒画素の密度を計算し、密度が $th_{density}$ (予め定めたしきい値) 以上のものを「べた塗り領域」と判定する。得られたべた塗り領域の集合に対して距離を尺度にしてクラスタリングを行ない、近接しているものをまとめて新たにべた塗り領域を構成する。このようにして得られたべた塗り領域の集合を $\{Paints\}$ と表す。

【0030】写真領域抽出部3では、 $\{Dots\}$ と $\{Paints\}$ の中で近接するものをまとめて写真領域の集合 $\{Photos\}$ を作る。次に、図形領域抽出部4の処理について説明する。

【0031】初期セグメンテーション部2において「巨大矩形」と判定されたものに対して、例えば、文献「信学論、J77-DII、1、pp91-100、(1994)」に記載された手法を用いて線分、円弧、円などの図形要素を抽出する。得られた図形要素の集合に対して距離を尺度にしてクラスタリングを行ない、近接するものをまとめて図形領域を抽出する。図形領域の集合を $\{Graphs\}$ とする。 $\{Graphs\}$ の各要素はそれらを内接する多角形で存在範囲が示されている。

【0032】次に、文字領域抽出部5の処理について説明する。ここでは、初期セグメンテーションで「文字候補矩形」と判定されたものに対して以下の処理を適用することにより文字領域を抽出する。すなわち、文字領域のレイアウト解析処理を行うことにより文字領域を抽出する。

【0033】文字領域のレイアウト解析処理とは、文字成分（文字候補矩形と判定され黒連結成分の外接矩形のサイズ、位置等の幾何的情報を含む）を最小単位として、これらの統合処理により、

- ・タイトル
- ・著者
- ・アブストラクト
- ・パラグラフ
- ・カラムで分離された本文

などの論理オブジェクトを抽出する処理である。さらには論理オブジェクトを構成する各領域において文字列を抽出する処理である。

【0034】図2を参照して文字領域抽出部5の処理（レイアウト解析処理）の概要を説明する。まず、領域抽出処理部10において、幾何情報に基づく処理を行う。領域抽出処理における下位階層である局所統合部10aでは、文字成分をあらかじめ設定した統合パラメータ（他の文字成分を統合するための距離範囲）に基づき統合する処理を行う。

【0035】領域抽出処理における上位階層である秩序形成部10bでは、局所統合部10aでの統合処理結果として2次元的な領域が生じたら、そこで

- ・文字列方向の判定
 - ・文字列の抽出
 - ・文字サイズ、文字間距離（字間）、行間距離（行間）の計算
- を行う。

【0036】そして、得られたパラメータ（これらを総称して秩序パラメータと呼ぶ）を次のように局所統合部10aで境界条件（統合パラメータ）として反映させる（局所統合部10aと秩序形成部10bの周縁制御）。

【0037】局所統合部10aで文字成分を統合する（局所統合）する際には、

- ・文字サイズおよび文字列方向が同じ部分領域間で統合を行う。
- ・文字列方向には字間分だけ、文字列方向と直行した方向には行間分だけ近接している部分領域を統合する（字間および行間値に基づいて統合パラメータが計算される）。

【0038】といった条件が満たされている必要がある。この処理サイクルを局所統合において新たな統合処理が生じなくなるまで繰り返す（孤立した文字すなわち領域を構成しない文字については統合パラメータを増加させる）。

【0039】この結果、統合範囲と統合対象が適応的に決定されながら部分領域が抽出され、文字サイズ・文字列方向・字間・行間が異なる領域が分離される。

【0040】次に、認識処理部11において、認識処理（意味情報に基づく処理）を行う。認識処理は領域抽出処理部10での領域抽出処理の上位階層に位置づけられており、その秩序形成の役割を持つ。ここでは、以下（ステップS1～ステップS4）のようにして認識処理結果により領域の秩序の形成を行なう。

- 10 【0041】ステップS1：文字認識部11aで各部分領域で文字列単位に文字認識処理を実施し、認識結果評価部11bで認識結果を評価して、さらに領域変更／棄却部11cで、
- ・非文字成分を検出して棄却する。

【0042】・認識結果の信頼度の低い文字列や部分領域に対してパラメータを変更する。といった選択処理を実施して処理の頑健性を高めるようにする。

ステップS2：正しく認識できたと判断された部分領域では秩序パラメータを抽出し、統合パラメータを再計算し、得られた値を自分の新たな境界条件とする（認識処理部11と領域抽出処理部10の間の周縁制御）。

【0043】ステップS3：各部分領域で「領域としてのまとまりの良さ」をその秩序性として評価する（これを部分領域の秩序度と呼び、秩序パラメータとは区別する）。低い秩序度を示す領域についてはその周辺の部分領域の秩序度を下げずに自らの秩序度を上げるように統合パラメータを制御して（境界条件として与えて）、領域抽出処理部10に対し、再び部分領域の統合（局所統合）を促す（部分領域の秩序度を高める）。

- 30 【0044】ステップS4：以上の領域抽出処理部10と認識処理部11との間の処理サイクルを全体の秩序度が上がらなくなるまで繰り返す。

次に、文字領域抽出部5におけるレイアウト解析処理について、図3～図5に示すフローチャートを参照して詳細に説明する。

【0045】まず、領域抽出処理部10の領域抽出処理について詳細に説明する。

ステップS10：局所統合部10aにおける文字成分の統合（部分領域の生成）処理（局所統合処理）

各文字成分には水平方向の統合パラメータhmpと垂直方向の統合パラメータvmpが付与されており、この値に基づいて水平方向と垂直方向に他の文字成分を統合して部分領域を生成する。

【0046】これらの統合パラメータの初期値は、例えば、あらかじめ「水平方向と垂直方向の文字成分間距離の最頻値」（文献「信学技報、PRU92-32、1992」参照）に基づいて、

$$hmp = h_s \times \theta$$

$$vmp = v_s \times \theta \dots (1)$$

- 50 と設定してもよいし、予め最小値として定めた値を付与

してもよい。

【0047】なお、

h s (v s) : 水平 (垂直) 方向文字成分間距離の最頻値

θ : (0 , 1) 内の定数

とする。

【0048】局所統合処理では以下の条件を満たすことを制約として課す。すなわち、

・統合によって生じた新たな領域がフィールドセパレータをまたがないこと。

・統合されるべき二つの部分領域では文字サイズがほぼ等しいこと。

【0049】・統合されるべき二つの部分領域では文字列方向が等しいこと。

*

$$S t r = \alpha \times 1 / C_0 + \beta \times 1 / L R C + \gamma \times (S_{p o t} \quad S_{s o} + S_{d o}) \quad \dots (2)$$

ここで、

C₀ : 文字並びの度合い (文献「信学技報、PRU92-32、1992」参照) 、

S_{pot} : 文字列の先頭位置の平均偏差

S_{so} : 文字列幅に関する平均偏差

S_{do} : 文字列間距離に関する標準偏差、

α 、 β 、 γ : 定数

L R C : 文字列方向の複雑度 (文献「PRU92-32」参照)

とする。

【0053】水平方向の文字らしさの尺度と垂直方向の文字列らしさの尺度を比較してその値が小さい方の文字列方向と文字列抽出結果を採用する。次に、文字サイズ・字間・行間値の推定を行う。すなわち、抽出された文字列から文字列の高さ・横書き文字列なら縦幅、縦書き文字列なら横幅)の平均値をその部分領域の平均的な文字サイズとし、文字列間距離を行間と見なしてそれぞれ抽出する。さらには得られた文字サイズを用いて各文字行内で文字サイズに満たない文字成分同士を統合して新たな文字成分を抽出し、さらに平均的な文字成分間距離を計算してその部分領域における平均的な字間と見なす。

【0054】ステップS12: 周縁制御による局所統合の規定

各部分領域で秩序形成部10bにおける秩序形成処理で得られたパラメータを局所統合に反映させるため、局所統合部10aでは、新たに得られた字間と行間値に基づいて水平方向と垂直方向の統合パラメータを以下の変更ルールに基づいて再設定して、局所統合を繰り返す。

【0055】変更ルール1 : 字間 (行間) 値が検出できた場合には、字間 (行間) + t h_aを新たな統合パラメータとして採用する。ここで、t h_aを予め定めたしきい値とする。

【0056】変更ルール2 : 字間 (行間) が検出されな

* ・統合されるべき二つの部分領域では字間および行間がほぼ等しいこと。

局所統合部10aにおける統合処理は、上記制限のもと、新たな統合が生じなくなるまで実施される。

【0050】ステップS11: 秩序形成部10bにおける部分領域の秩序の形成

局所統合部10aでの局所統合が安定した時点で以下の処理を適用して各部分領域で秩序を形成する。

【0051】まず、文字列方向の検出および文字列の抽出を行う。すなわち、局所統合部10aで生成された各部分領域について、水平方向と垂直方向に文字列を抽出して、両方向で以下に定義する文字列らしさを表す尺度S t rを計算する。

*

【0052】

かった場合には、既に設定されている統合パラメータを予め設定した増し分oだけ増加させて新しい統合パラメータとする。

20 【0057】ステップS13: 以上、ステップS10〜ステップS12の局所統合処理、秩序形成処理、周縁制御の処理サイクルを新たな部分領域の統合が生じなくなるまで繰り返し、その結果得られた部分領域に対して、認識処理部11において認識処理を行う。

【0058】次に、認識処理部11の認識処理について詳細に説明する。ここでは、前述の領域抽出処理部10のレイアウト解析による領域抽出処理により得られた部分領域とそれを構成する文字列に対して以下の手順で認識、評価、選択、変更を行う。

30 【0059】ステップS20: 文字認識

まず、文字認識部11aにおいて、領域抽出部10で抽出された各部分領域の各文字列に対して、文献「信学技報PRU93-47、(1993-09)」に記載された手法に基づいて文字切り出し/認識処理を実施する。この処理では分離文字が統合され、また接触文字が切断されるために正しい文字サイズが得られる。

【0060】ステップS21: 認識結果の評価

続いて、認識結果評価部11bにおいて、各部分領域で文字認識結果 (類似度) の平均値 r c g_{ave} を計算する。

40 r c g_{ave} が予め設定されているしきい値 t h_{rec} より低い部分領域は低信頼度の部分領域、高い場合には高信頼度の部分領域と判定する。

【0061】ステップS22: 認識結果に基づく部分領域の変更/棄却処理

さらに、領域変更/棄却部11cにおいて、信頼度に基づいて部分領域に対して次の変更/棄却処理を適用する。

【0062】ステップS22a: 低信頼度の部分領域に対する処理 (文字サイズを変更させて新たな部分領域を発生させる)

50

・部分領域内で文字サイズに関する出現頻度のヒストグラムを計算する。

【0063】・このヒストグラムに複数のピークが存在する場合には、それに基づいて文字サイズ情報を変更して、領域抽出処理部10で前述の領域抽出処理をやり直す。

・ヒストグラムに現状以外のピークが存在しない場合には、着目文字列を棄却する。

【0064】ステップS22b：一方、高信頼度の部分*

$$\text{秩序度} = \frac{\text{部分領域の文字行数}}{\min(\text{最大文字行数を持つ部分領域の文字行数}, \gamma)}$$

・・・(3)

【0066】ここで、

$\min(A, B)$ ：AとBのうち小さい方を選ぶ関数
 γ ：予め設定した値とする。

【0067】秩序度が予め定めたしきい値 $t_{h_{order}}$ 以下の部分領域を低秩序度の部分領域、 $t_{h_{order}}$ をこえる部分領域を高秩序度の部分領域と見なし、高秩序度の部分領域と、低秩序度の部分領域のそれぞれについて以下の処理を行う。

※

$$\text{垂直(水平)方向への統合範囲の限界} = \begin{cases} \lambda & (\text{近接する部分領域が存在する場合}) \\ \mu & (\text{それ以外}) \end{cases}$$

・・・(4)

【0070】ここで、

$$\lambda = \min_{ad} / d_1$$

\min_{ad} ：最近の部分領域との距離

d_1 ：予め定めた値

μ ：予め設定した定数

とする。

【0071】統合範囲が設定されている部分領域では統合範囲以上に離れている部分領域とは統合しないこととする。

ステップS22b-2：低秩序度の部分領域の統合パラメータの変更

低秩序度の部分領域に対して、他の部分領域と統合し安くする(自分の秩序度を上げる)ことを目的として水平方向と垂直方向の統合パラメータ(h_{mp} 、 v_{mp})を次式により変更する。

【0072】

$$\text{統合パラメータ} = \text{統合パラメータ} \times \delta \quad \dots (5)$$

ここで、 δ は1より大きい値をとり徐々に増大していく関数であるとする。文字間隔が不規則でスパースな部分領域で統合が促進されるようにする。

【0073】ステップS22b-1、ステップS22b-2での処理で変更された秩序パラメータおよび統合パラメータを下位の領域抽出処理部10に境界条件として渡し、下位レベルの処理を規定する(領域抽出処理部1

*領域に対しては、秩序パラメータを計算し、統合パラメータを変更する。まず、各部分領域に対して以下に定義する式に基づいて秩序度を計算し、「文字成分が密集している安定したもの」と「少ない文字成分により構成されている不安定なもの」に分類する。秩序度は次式(3)により決定される。

【0065】

【数1】

※【0068】ステップS22b-1：高秩序度の部分領域の統合範囲の設定

高秩序度と判定された部分領域では隣接する(最近の)高秩序度部分領域との距離に基づき次式(4)で定義される自分の統合範囲の限界が設定される(統合範囲の抑制)。

【0069】

【数2】

$$\begin{cases} \lambda & (\text{近接する部分領域が存在する場合}) \\ \mu & (\text{それ以外}) \end{cases}$$

・・・(4)

0と認識処理部11間の周縁制御)。

【0074】各領域に秩序度が付与されたあとの領域抽出処理部10における局所統合は次のように規定される。

・統合により新たに生じる部分領域は他の部分領域と重ならない

・各部分領域で文字列らしさの尺度が統合前と比べて低下しない

・各部分領域で秩序度が統合前と比べて低下しない

・統合が衝突する場合にはそれらの中で最良のものを選択する

このような規定のもと、あらたに設定された統合パラメータに基づく統合範囲内に存在し、かつ、文字列方向が同じで、文字サイズが類似した他の部分領域を統合していく。

【0075】ステップS23：以上ステップS10～ステップS22bの領域抽出処理部10と認識処理部11との間の処理サイクルを全体の秩序度、すなわち、(3)式から計算される秩序度が上がらなくなるまで繰り返した結果得られた文字領域の集合を{Text}と表す。

【0076】次に、文字領域抽出部5のレイアウト解析処理の第2の実施形態について説明する。この実施形態は、自律分散的な処理形態であることを特徴とする。す

なわち、領域抽出処理における各部分領域に対して3つのプロセスを与える。プロセスの内訳は以下のとおりである。

【0077】・統合プロセス(自分で持つ統合範囲内に近接する他の部分領域を統合する)

・幾何的秩序形成プロセス(自分の領域の幾何的な秩序性(文字列方向・文字列・文字サイズ・字間・行間)を抽出する)

・意味的秩序形成プロセス(自分の領域内の文字列を認識する)

これらのプロセスは、図6に示すように階層構造を有した組み(ユニット)となっている。この場合、計算空間内で部分領域の数だけユニットが存在する。

【0078】ユニット内では統合プロセスと幾何的秩序形成プロセス間、および、幾何的秩序形成プロセスと意味的秩序形成プロセス間は同期していない。各上位プロセスはその下位プロセスで情報が抽出され次第動作するようになっている。

【0079】各プロセスの動作は次のとおりである。

・統合プロセス

他の統合プロセスを統合する(統合に関する判断は前述同様)。2つの統合プロセス間で統合が生じたら優勢な情報(より多くの情報)を有している統合プロセスの方に融合され一つとなる(統合したプロセスが有する情報を吸収してユニットごと殺す)。さらには、自分の統合範囲および統合相手の適正を判断する。

【0080】・幾何的秩序形成プロセス

自分の領域の幾何的な秩序性(文字列方向・文字列・文字サイズ・字間・行間)を抽出して、統合プロセスの運動を規定する(規定の仕方は前述同様)。

【0081】・意味的秩序形成プロセス

自分の領域内の文字を認識して、不要な情報を棄却したり、変更する。さらに、下位の幾何的秩序形成プロセスを規定する(規定の仕方は前述同様)。

【0082】このようなユニット群はユニット全体の秩序性が安定する方向に動く。これは例えば、共有メモリを別途設けておき、ここに各ユニットの処理結果である部分領域およびその秩序パラメータが書き込まれるようにする。この共有メモリ上では情報の書き込み、書き換え、消去が各ユニットにより行なわれる。情報が書き込まれたら共有メモリ上で部分領域の分布に対する秩序度が共有メモリ自身によって計算される(内容は前述同様)。

【0083】各ユニットは共有メモリを観察し、秩序度の分布に応じて(秩序度が上昇するように)ユニットの動作(特に統合ユニットにおける統合処理すなわち自分はその相手と統合するか)を決める。

【0084】この自律的な処理が安定したところでレイアウト解析処理を終了する。以上の処理で得られた文字領域の集合を{Text}と表す。文字領域抽出部5の

第3の実施形態について説明する。

【0085】前述の文字領域抽出部5のレイアウト解析処理の領域抽出処理における文字列の抽出処理は以下の処理により実現されていてもよい。すなわち、この処理では各部分領域で、文字成分から、文字の大きさ、並び、文字列方向が同じ物をグループ化して文字列として抽出するものである。

【0086】まず、文字候補矩形に関する情報の抽出を行う。文字成分に対してそれを包含する最小の矩形を定義する。図7に示すように、重なりあっている矩形に対しては、それらを包含する最小の外接矩形Gを定義し、これを文字矩形と呼ぶ。この文字矩形に対しては縦横比が求められる。各矩形内部における画像の性質に対してストロークの方向成分の分布、ストロークの太さ、複雑度が求められる。以後それらを統合した情報を矩形情報と呼ぶ。

【0087】次に、この矩形情報から隣接関係に関する情報の抽出を行う。すなわち、図8に示すように、各外接矩形について隣接している外接矩形同士の間隣接関係を定義する。その手法としては、例えば、隣接関係によって結ばれる矩形間の間隔や複数の隣接関係の整列の程度を求める。整列の程度としては、例えば、図9に示すように、矩形の重心を結ぶベクトルや水平、垂直方向の重なりを用いる。これらの評価値を統合し、矩形同士のつながりの強さ(以下接続強度とよぶ)を求める。

【0088】評価値の統合の方法は例えば単に隣の矩形に対する評価値について、予め定義された重みで各評価値の線形和を取り、それを接続強度としてもよいし、一旦隣り合うもの同士の接続強度を求めた後、前後数個の中で着目矩形に近い程重視するように重み付けし、畳み込んでよい。この様に定義された接続強度は各々の外接矩形が同一文字列に属す確信度を示す。また、これを部分矩形列で統合したものが部分矩形列の確信度を示す。

【0089】この矩形列は、図10に示すように、ノードに矩形情報を有し、エッジが或る接続強度で重み付けされたグラフで表現でき、以後これを隣接関係グラフと呼ぶ。

【0090】次に、部分隣接関係グラフを抽出する。すなわち、文字矩形に関する情報、隣接関係に関する情報の双方を用いて、文字列に相当する矩形の並び(部分グラフ)を抽出する。図11を参照して具体的に説明すると、まず、隣接関係グラフに対して部分領域中から統計的に求められた一定値(接続強度)より強い隣接関係のみを選択的に抽出することにより部分隣接関係グラフの初期状態を得る(図11(a)参照)。そして、部分隣接関係グラフの初期状態をもちいて統計的に求められた文字矩形に関する情報、隣接関係に関する情報を用いて、部分隣接関係グラフの理想的状態を得る(図11(b))。また、そのグラフの統計量から予想される仮

想文字矩形を発生させ、初期状態から次の理想的状態を得る(図11(c))。さらに、理想的隣接関係グラフの状態と現実の部分領域とのギャップを文字矩形、隣接関係の各々に関する尺度で計算する。その際に矩形の統合、分離が発生する。このギャップが一定値を越えた時、理想的な部分隣接関係グラフに、それ以後新たな仮想文字矩形は生成されない。以上の処理を定常状態になるまで繰り返す。

【0091】さて、これまでに求められた部分隣接グラフ中で同一の矩形について、2種類以上の解釈が存在する場合、それを以後の処理に文字列候補として保存しておく必要がある。複数の解釈が存在する場合、一つの解釈にのみ基づく仮想文字矩形を生成し、その他の解釈を排除する事によって、対象となる隣接関係グラフの全ての解釈の文字列候補を生成することが可能である。

【0092】仮想的な文字矩形として、例えば次のような場合がある。すなわち、図12に示すように、水平方向と垂直方向の部分隣接関係グラフに同時に含まれている場合、図13に示すように、行の構成上、2行の部分隣接関係グラフと1行の隣接関係グラフのいずれかに含まれている場合等である。

【0093】また、このように生成された各仮説における評価値と、分岐点付近における接続強度によって、仮説を少数に限定することもできる。ところで、入力されてくる文書の方向が未知である場合には、その方向を把握する必要がある。この場合、前述の認識処理部11において、上下左右の4方向の可能性を考慮して文字認識処理を行わなければならない。このときすべての文字列に対して4方向で文字認識を行なうことが考えられるが、処理量は文字方向が既知である場合の4倍となり効率的でない。そこで、文書全体において、または、部分領域ごとにある一つの文字列を選び、それに対してその文字方向を特定することによって、その情報に基づいて残りの文字列を効率的に読み取ることが可能となる。

【0094】この場合、一つの文書あるいは部分領域中では文字の向きは同一であると仮定し、文字領域中から一行の文字列を抽出し、この文字列に対してのみ4方向で文字認識を行ない、その中から最も認識率の良い方向を選び、その方向で残りの文字列を読み取るようにしてもよい。

【0095】次に、文字方向決定処理の具体例について説明する。

ステップS30：これまでに求められた文字列候補のうち、文字矩形数が最大のものを選ぶ。

【0096】ステップS31：ステップS30で選択された文字列候補を4方向で認識する。例えば文字列パターンから90度、180度回転させたパターンを生じさせ、これらを前述の文字認識手法によって認識させる。

【0097】ステップS32：ステップS31で得られた認識結果のうちで平均文字認識率が最も良い方向を選

択する。

ステップS33：ステップS32で得られた認識結果が有効であるか否かについて判定する。この場合、例えば平均文字認識率がしきい値TH1より大きい場合を「有効」、それ以外を「無効」と判定する。

【0098】ステップS34：ステップS33の判定結果が「無効」である場合には無効となった文字列以外の文字列に対してステップS30からステップS33までを繰り返す。なお、「有効」である場合には処理を終了する。

【0099】この処理結果をもとに、文書画像あるいは部分領域の処理結果を得られた角度方向を解消する方向に回転するようにしてもよい。これにより以後の認識処理は必ず正しい方向で行なわれる。

【0100】以上、写真領域抽出部3、図形領域抽出部4、文字領域抽出部5における処理が終了した結果、どの部分領域にも成りえなかった画像オブジェクトに対して「その他」という属性を付与する。これは後段の処理において文字、写真、図形の何れかの属性をもつものとし、この段階では「不明」とする。

【0101】また、この時点で、
・文書画像の大部分が{Dots}で占められている
(文書画像が全体的にかすけている可能性があるとしてオペレータに再入力促す)

・文書画像の大部分が{Paints}で占められている
(文書画像が全体的につぶれている可能性があるとしてオペレータに再入力促す)
という結果が得られている場合には、文書画像の再入力をオペレータに促してもよい。

【0102】以上の処理により、

- ・写真領域の集合：{Photos}
- ・図形領域の集合：{Graphs}
- ・文字領域の集合：{Texts}

が得られている。

【0103】この時点では、図14に示すように、同一の画像オブジェクトが複数種類の領域に属している(領域間で重複がある)可能性がある。一般に文書では異種領域間で重なりが生じない(排他的である)としているので、部分領域間で重複している場合にはそれらを解消する必要がある。

【0104】また、図の記述に用いられている文字列や、写真中に存在する文字なども文字成分として抽出されている可能性がある。この場合、例えば、図中の文字列は本文を構成する文字列と区別され(図形領域と文字領域の排他的関係)、写真中の文字は写真扱いにするほうが好ましい(写真領域と文字領域の共存関係)。このような領域間の排他的関係、共存関係に関する情報は、あらかじめ領域間重複解消部6に記憶され、管理されている。これを用いて領域間重複解消部6では、各領域抽出部3、4、5で抽出された領域の重なりを解消する処

理を行うようになっている。

【0105】そこで、次に、図1の領域間重複解消部6において実行される領域の重なりを解消する処理と、図形領域中の文字列を識別する処理について述べる。まず、写真領域の集合 {Photos}、図形領域の集合 {Graphs}、文字領域の集合 {Texts} をそれぞれ部分集合とする全体集合 {Area

cand} を生成する。

*

着目部分領域の面積

$$B_p = \frac{\text{着目部分領域の面積}}{\max(\text{文書中の最大の文字領域の面積、文書中の最大の写真領域の面積})} \dots (6)$$

【0108】ステップS41：図形領域に対する確信度 B_g の付与

部分領域のサイズに基づいて、次式(7)に従って図形※

※領域に確信度 B_p を付与する。

【0109】

【数4】

着目部分領域の面積

$$B_g = \frac{\text{着目部分領域の面積}}{\max(\text{文書中の最大の文字領域の面積、文書中の最大の図形領域の面積})} \dots (7)$$

【0110】ステップS42：文字領域に対する確信度 B_t の付与

式(3)により計算される秩序度を確信度 B_t として採用する。

ステップS43：複数の属性を有する部分領域の確信度の再計算

★まず、{Area_{cand}} の要素のうち、他の要素と領域が重なるものを複数の属性を有する部分領域として抽出する。そして、複数の属性を有する部分領域の確信度 B_{mix} を次式(8)に基づいて再計算する。

【0111】

【数5】

★着目部分領域の面積

$$B_{mix} = \frac{\text{着目部分領域の面積}}{\text{重なっている部分領域をANDすることにより得られる面積}} \dots (8)$$

【0112】次に、{Area_{cand}} に対して、

- ・領域は重ならない
- ・領域は文書画像の構成要素の全てを被覆する

という幾何的条件に基づいて、同時に共存可能な領域候補の組み合わせの集合 {Area_{comb}} を求める。

【0113】同時に共存可能な領域候補の組み合わせは、例えば、文献「信学技報、PRU94-32、1994」に記載されている連合グラフ法を用いて、以下の手順により作成されてもよい。この方式を用いることにより同時に成立できる領域候補の組をグラフを用いて表現することができ、クリーク抽出というグラフ理論的手法により排他的な領域候補の組み合わせの集合を正確に抽出することが可能となる。

【0114】ステップS50：{Area_{cand}} における各領域候補を連合グラフのノードとして割り当てる (図15参照)。

ステップS51：{Area_{cand}} における全ての2つの割り当てにおいてそれらが両立するか否かを判定する。2つの領域候補に重なりがない場合にはその2つの領域は両立すると見なす。

* 【0106】そして、以下の手順で各領域に確信度が付与される。

ステップS40：写真領域に対する確信度 B_p

p の付与部分領域のサイズに基づいて、次式(6)に従って写真領域に確信度 B_p を付与する。

【0107】

【数3】

着目部分領域の面積

$$B_p = \frac{\text{着目部分領域の面積}}{\max(\text{文書中の最大の文字領域の面積、文書中の最大の写真領域の面積})} \dots (6)$$

【0108】ステップS41：図形領域に対する確信度 B_g の付与

部分領域のサイズに基づいて、次式(7)に従って図形※

※領域に確信度 B_p を付与する。

【0109】

【数4】

着目部分領域の面積

$$B_g = \frac{\text{着目部分領域の面積}}{\max(\text{文書中の最大の文字領域の面積、文書中の最大の図形領域の面積})} \dots (7)$$

【0110】ステップS42：文字領域に対する確信度 B_t の付与

式(3)により計算される秩序度を確信度 B_t として採用する。

ステップS43：複数の属性を有する部分領域の確信度の再計算

★まず、{Area_{cand}} の要素のうち、他の要素と領域が重なるものを複数の属性を有する部分領域として抽出する。そして、複数の属性を有する部分領域の確信度 B_{mix} を次式(8)に基づいて再計算する。

【0111】

【数5】

★着目部分領域の面積

$$B_{mix} = \frac{\text{着目部分領域の面積}}{\text{重なっている部分領域をANDすることにより得られる面積}} \dots (8)$$

【0112】次に、{Area_{cand}} に対して、

- ・領域は重ならない
- ・領域は文書画像の構成要素の全てを被覆する

という幾何的条件に基づいて、同時に共存可能な領域候補の組み合わせの集合 {Area_{comb}} を求める。

【0113】同時に共存可能な領域候補の組み合わせは、例えば、文献「信学技報、PRU94-32、1994」に記載されている連合グラフ法を用いて、以下の手順により作成されてもよい。この方式を用いることにより同時に成立できる領域候補の組をグラフを用いて表現することができ、クリーク抽出というグラフ理論的手法により排他的な領域候補の組み合わせの集合を正確に抽出することが可能となる。

【0114】ステップS50：{Area_{cand}} における各領域候補を連合グラフのノードとして割り当てる (図15参照)。

ステップS51：{Area_{cand}} における全ての2つの割り当てにおいてそれらが両立するか否かを判定する。2つの領域候補に重なりがない場合にはその2つの領域は両立すると見なす。

【0115】ステップS52：両立する割り当て間にエッジを設定することにより連合グラフを作成する (図16参照)。

ステップS53：連合グラフからすべての極大クリークを抽出する (図16参照)。極大クリーク抽出手順は、例えば文献「信学論(D)、J68-D、3、pp221-228、(1985)」に記載されている手法を用いることにより抽出される。図16では、ノード3個からなる共存関係の領域候補の組み合わせとしての極大クリークを抽出している。

【0116】ステップS54：極大クリークのうち文書画像のすべての構成要素を被覆しているものを抽出する。得られた極大クリークの集合は同時に共存可能な領域候補の組み合わせ集合 {Area_{comb}} とする。

【0117】この手順の他に、各部分領域を仮説と見なし、部分領域の組み合わせを仮説の組み合わせとしてATMS (Assumption based Truth Maintenance System) を用いて仮説間の無矛盾性を管理することにより共存可能な領域候補の組み合わせを得るようにしてもよい。

【0118】最終的な領域分割結果は、領域候補の組み合わせ集合 $\{Area_{comb}\}$ の中から最良な組み合わせを抽出することで得られる。これは例えば、図16のグラフ表現された組み合わせ集合に対して(9)式に示す評価関数に基づいて、極大クリークの各エッジにコスト*

$$Cost_{edge} = \phi \times Dist \times Posi$$

* $Cost_{edge}$ を付与し、各極大クリークごとに総コストを計算し、最もコストの低い極大クリーク(領域候補の組み合わせ)を抽出することによって達成される。

【0119】

【数6】

... (9)

$Cost_{edge}$: エッジのコスト

$$\phi = \begin{cases} \beta & (\text{二つの領域が同質であるとき}) \\ \gamma & (\text{二つの領域が異質であるとき}) \end{cases}$$

$Dist$: 領域間の距離

$$Posi = \begin{cases} \delta & (\text{二つの領域が上下左右に重なっているとき}) \\ \epsilon & (\text{それ以外のとき}) \end{cases}$$

この評価関数によって極大クリークの各エッジにコスト ※【0120】
が割り振られたら、例えば、 ※【数7】

$$Cost_{all} = \frac{\sum Cost_{edge} \times BF}{\text{エッジの総数}} \quad \dots (10)$$

【0121】といった評価関数を用いて各極大グラフで総コスト $Cost_{all}$ を計算し、その値が最小となる極大グラフを抽出し、それが表す部分領域の組み合わせを★

★領域分割結果と見なしてもよい。

【0122】ここで、

$$BF = 1 / (\text{エッジの両端の部分領域の確信度の積}) \quad \dots (11)$$

とする。

することにより実施される。

【0123】このとき、コスト値の昇順に複数個の排他的な領域候補を出力することも可能である。このような場合、正しい候補をオペレータが指定できるように複数候補を画面に出力するようにしてもよい。

【0128】規則：図形領域の属性を持つ部分領域においてその内部の線分のほとんどが水平線か垂直線のいずれかであり、内部には高信頼度の文字列が含まれており、それらが線分で囲まれているならば、その部分領域を表領域と見なす。

【0124】さらに、最適な組み合わせを抽出する方法として、

【0129】この規則により、表として認識された部分領域に対しては「表」を意味する属性を新たに付与する。また、表を除く図形領域中の部分領域に対しては、文字領域の部分領域の抽出時に得られた情報を用いて、当該部分領域に含まれる文字列情報(認識結果であるコード情報を含む)を無効とせず、それらに対して「図形を説明する文字列」という属性を付与しておく。これらの文字列は文書検索時に活用されるようにしておいてもよい。

・各部分領域に属性に応じて確信度を付与し、部分領域間の局所的な整合性を計算し、弛緩法などを用いて各部分領域の確信度を反復的に更新していくことにより最終的に信頼度の高い解を得る。

【0125】・画像全体に対して領域の整合性を表すエネルギー関数を設定し、関数をシミュレーテッドアニーリング法(統計的緩和法)などを用いてこのエネルギー関数を最小にする組み合わせを探索する。

【0126】などの方法を用いて求めるようにしてもよい。領域間重複解消部6における上述した領域分割処理の結果、領域候補間のあいまいさが解消される。この結果、確定されなかった領域候補を無効とする。確定された領域では各領域は均質であるとして、その内部に含まれるすべての画像オブジェクトに対してその領域の属性を新たに付与する。

【0130】さらに、最終的に得られた領域分割結果において、各領域は例えば図17に示すように公知の方法により両域内の画像を含む最小の凸な多角形で囲み、それをその領域の範囲としても良い。各領域ではそれを構成する部分領域と画像オブジェクトに関する情報が、画像→領域→部分領域の集合→オブジェクトの集合というように階層的に管理されていてもよい。

【0127】さて、図形領域と確定された領域では、さらにグラフ、表、図などの種類が考えられる。表の場合はその中の文字列が重要な情報となるのでグラフや図と区別する必要がある。表の識別は例えば次の規則を適用

【0131】以上、説明したように、上記実施形態によれば、画像入力部1で入力された文書画像に対しラベリング処理を行って黒連結成分を抽出して、その外接矩形のサイズから写真領域抽出部3、図形領域抽出部4、文

字領域抽出部 5 のいずれかに振り分け、写真領域抽出部 3 では、べた塗り領域、ドット領域の抽出処理を行い、図形領域抽出部 4 では、例えば、線分、円弧、円などの図形要素を抽出して、得られた図形要素の集合に対して距離を尺度にしてクラスタリングを行ない、近接するものまとめて図形領域を抽出する処理を行い、さらに文字領域抽出部 5 では、統合パラメータに基づき初期セグメンテーション部 2 で抽出された文字成分あるいは文字成分を統合して得られた部分領域を統合する局所統合処理を行い、この統合された部分領域に対し、文字列らしさの評価を行って文字列を抽出するとともに、その文字列の幾何的な特徴量（秩序パラメータ）を求め、この秩序パラメータをもとに統合パラメータを再設定しながら部分領域の統合がなくなり安定するまで部分領域の統合を行い、さらに、その統合された部分領域を構成する文字列の文字認識を行い、その認識結果から信頼度の高い部分領域について、秩序度を算出し、その秩序度が上がるように、統合パラメータを変更しながら部分領域を統合して前記処理を繰り返して行うことにより文字領域を抽出し、写真領域抽出部 3、図形領域抽出部 4、文字領域抽出部 5 のそれぞれで抽出された領域間に空間的重なりが見られる場合には、領域重複解消部 6 で各領域間の排他関係と共存関係に基づき、例えば、グラフ理論的手法により最適な組み合わせを抽出することにより、書式が未知であり、文書の入力方向が未知である文書画像において、文字領域、写真領域、図形領域等の性質の異なる複数の画像領域を高精度に分離・識別できる。

【0132】また、文字領域抽出部 5 では、文字成分、部分領域といった小領域単位に幾何的秩序、意味的秩序の解析を行うことにより、高精度な文字領域の識別が行える。

【0133】さらに、文章、写真、絵、図形、グラフ、表などの種々の成分から構成される様々な書式を持つ文書においてそれぞれの成分を高精度に分離・識別することができ、図形、グラフ、表などではそれらに含まれる文字成分を抽出することができるので、

- ・文章成分のみ利用したい
- ・タイトルやパラグラフ単位に適切にブロック化された文章成分を抽出したい
- ・図形成分のみ利用したい
- ・写真成分のみ利用したい
- ・図形内に記載されているキーワードを用いて図形を検索したい

といった様々なアプリケーションの高度な要求に答えることが可能となる。

【0134】

【発明の効果】以上説明したように本発明によれば、書式が未知であり、文書の入力方向が未知である文書画像において、文字領域、写真領域、図形領域等の性質の異なる複数の画像領域を高精度に識別できる画像処理方法

および画像処理装置を提供できる。

【図面の簡単な説明】

【図 1】本発明の一実施形態に係る画像処理装置の構成を概略的に示したブロック図。

【図 2】図 1 の文字領域抽出部の構成を概略的に示したブロック図。

【図 3】文字領域抽出部の処理動作の一具体例を説明するためのフローチャート。

【図 4】文字領域抽出部の処理動作の一具体例を説明するためのフローチャート。

【図 5】文字領域抽出部の処理動作の一具体例を説明するためのフローチャート。

【図 6】文字領域抽出部の第 2 の実施形態について説明するための図で、自立分散的な処理形態の一具体例を示した図。

【図 7】文字領域抽出部の第 3 の実施形態について説明するための図で、文字矩形の一具体例を示した図。

【図 8】文字領域抽出部の第 3 の実施形態について説明するための図で、文字矩形間の隣接関係を説明するための図。

【図 9】文字領域抽出部の第 3 の実施形態について説明するための図で、文字矩形間の接続強度を説明するための図。

【図 10】文字領域抽出部の第 3 の実施形態について説明するための図で、文字矩形の集合を隣接関係グラフで表した一具体例を示した図。

【図 11】文字領域抽出部の第 3 の実施形態について説明するための図で、隣接関係グラフから接続強度の強い部分隣接関係グラフを抽出して、理想的な部分隣接関係グラフを求めながら文字矩形の統合、分離を行う手順について説明するための図。

【図 12】文字領域抽出部の第 3 の実施形態について説明するための図で、部分隣接関係グラフの複数の解釈の一具体例について説明するための図。

【図 13】文字領域抽出部の第 3 の実施形態について説明するための図で、部分隣接関係グラフの複数の解釈の他の具体例について説明するための図。

【図 14】写真領域抽出部、図形領域抽出部、文字領域抽出部における処理が終了した時点で、処理対象文書の画像から抽出された領域の一具体例を示した図で、1つの画像オブジェクトが複数の領域に属する場合について説明するための図。

【図 15】図 1 の領域間重複解消部において、グラフ理論的手法により排他的な領域候補の組み合わせの集合を抽出する処理を説明するための図で、連合グラフを構成するノードの一具体例を示した図。

【図 16】連合グラフとそこから得られる極大クリークの一具体例を示した図。

【図 17】図 1 の領域間重複解消部での領域分割の結果、最終的に得られた領域の範囲の表現の一具体例を示

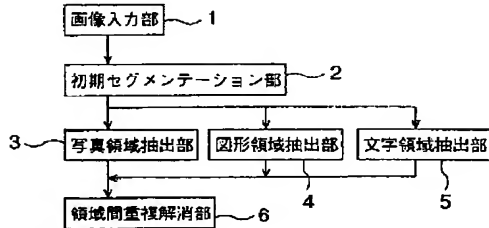
した図。

【符号の説明】

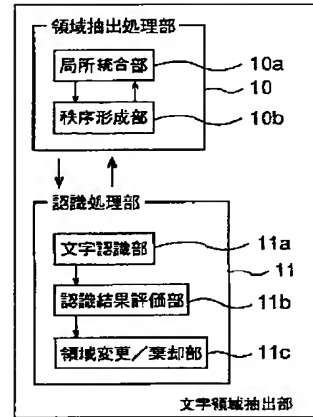
1…画像入力部、2…初期セグメンテーション、3…写*

* 真領域抽出部、4…図形領域抽出部、5…文字領域抽出部、6…領域間重複解消部。

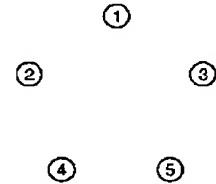
【図1】



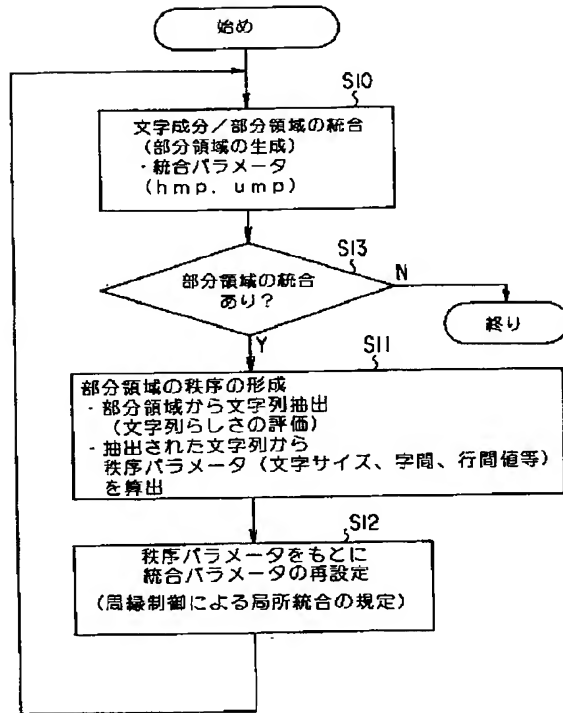
【図2】



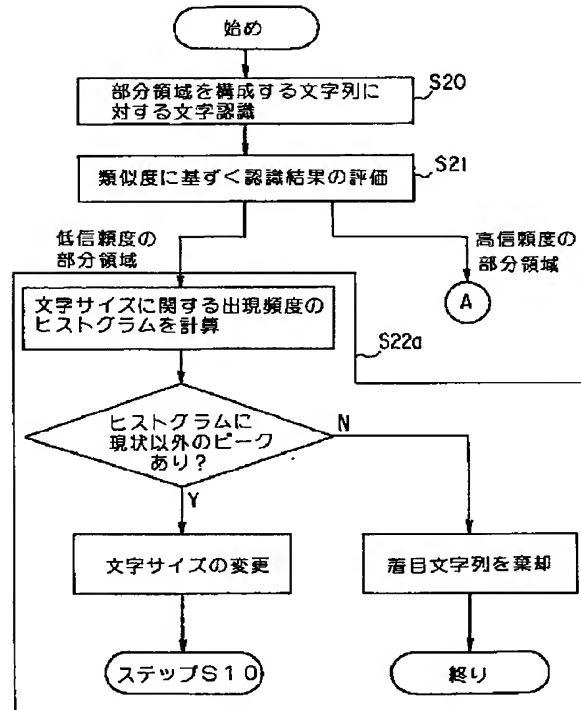
【図15】



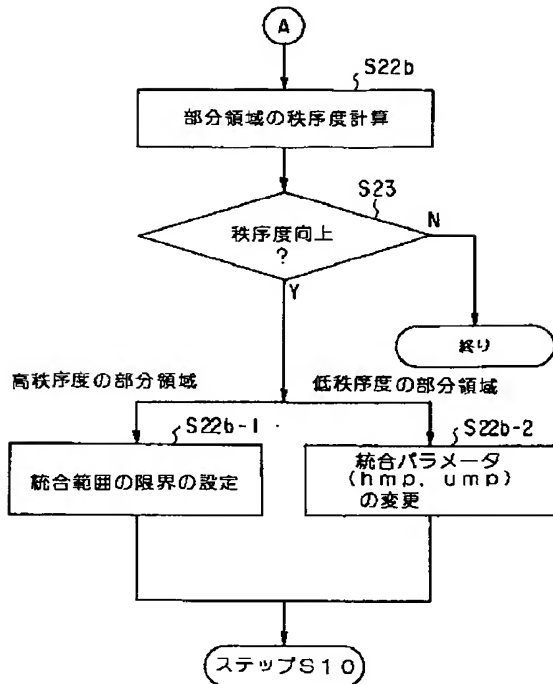
【図3】



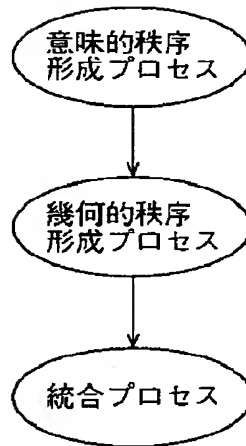
【図4】



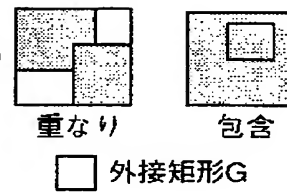
【図5】



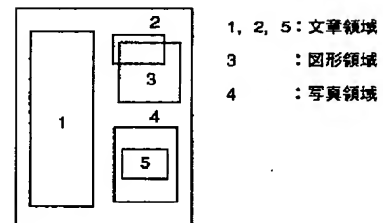
【図6】



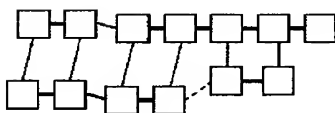
【図7】



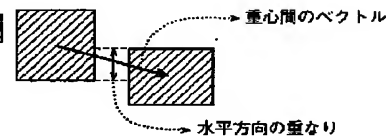
【図14】



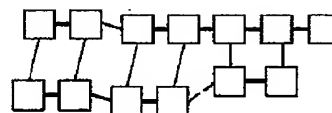
【図8】



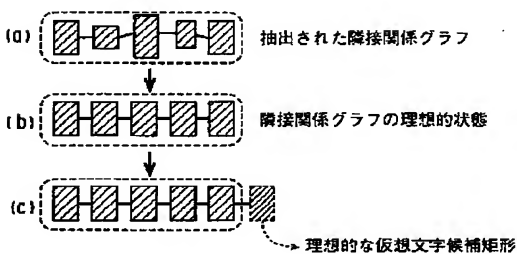
【図9】



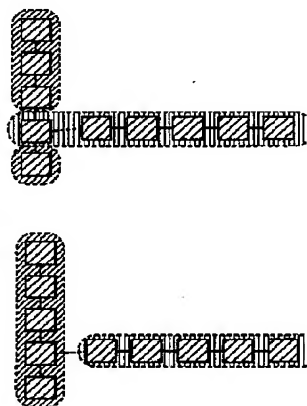
【図10】



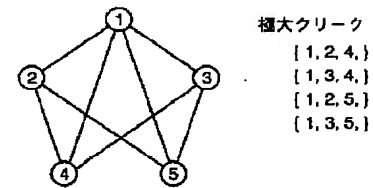
【図11】



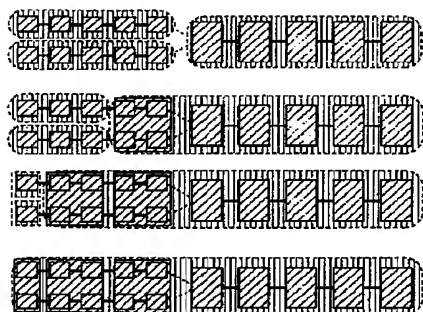
【図12】



【図16】



【図 13】



【図 17】

